File Number: 25-13

Joint
Computer-aided Acquisition
and Logistic Support (JCALS)
CALS Technology Center (CTC)

APPLICATION OF COMPUTER-ASSISTED DATA ACCEPTANCE (CADA) TECHNIQUES TO JOINT ENGINEERING DATA MANAGEMENT INFORMATION and CONTROL SYSTEM (JEDMICS)

FUNCTIONAL DESCRIPTION

CONTRACT NO.: DAAB07-93-D-T001
TASK NO.: 93-023

30 November 1995

Prepared for:

Department of Defense
PM JEDMICS

Prepared By:

ACCURATE Information Systems, Inc.
Meridian Center 1
2 Industrial Way West
Eatontown, New Jersey 07724

APPLICATION OF COMPUTER-ASSISTED DATA ACCEPTANCE (CADA) TECHNIQUES TO JOINT ENGINEERING DATA MANAGEMENT INFORMATION and CONTROL SYSTEM (JEDMICS)

FUNCTIONAL DESCRIPTION

CONTRACT NO.: DAAB07-93-D-T001
TASK NO.: 93-023

30 November 1995

Addendum

15 November 1996

Prepared for:

Naval Supply Systems Command
Joint Engineering Data Management Information
and Control System (JEDMICS)

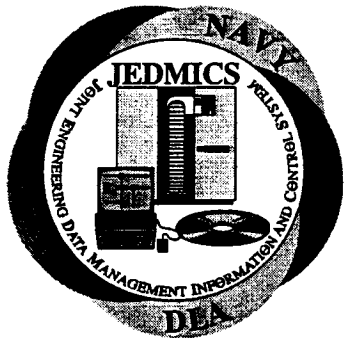

TABLE OF CONTENTS

LIST OF FIGURES

DISCLAIMER

The use of trade names in this document and/or discussion of a particular product does not constitute an endorsement or approval of the use of such commercial equipment. This document may not be cited for the purpose of advertisement.

The use of COTS products does not constitute an endorsement or approval of their use This software may not be cited for the purpose of advertisement. Any of the products listed in Appendix A may be replaced by equivalent products by the Government since they own the CADA source code.

SECTION 1 GENERAL

1.1 Purpose of the Functional Description

The objective of the Functional Description for the application of Computer-Assisted Data Acceptance (CADA) tools to Joint Engineering Data Management Information and Control System (JEDMICS) is to provide the reader with a high level description of the interface and system requirements. The Functional Description will serve as the basis for mutual understanding of interface requirements between JEDMICS and the CADA Quality Assurance (QA) platform. All enhancements made to the existing CADA tools, as requested by JEDMICS, are included in this Functional Description (FD).

1.2 Project References

The following references include the project contract and delivery order, contractor deliverables, applicable standards, and proprietary vendor documentation.

1.2.1 The Project Contract and Delivery Order

Contract No. DAAB07-XX-D-T001
Task Assignment Plan XX-023

1.2.2 Contractor Deliverables

Title	Date
CADA End Users Manual	*28 October 1994*
JEDMICS/CADA Contractor	*28 October 1994*
Final Report: Application of CADA to JEDMICS	*28 October 1994*

1.2.3 Applicable Standards

1.2.4 Proprietary Products

The following Commercial-Off-The-Shelf (COTS) products are integrated into the CADA tools, and licenses are required to use these software packages (See Appendix A).

Vendor Product[1]	Usage
NestorReader™	Intelligent Character Recognition (ICR) Product
ScanFix™	TMS Sequoia Software Libraries for line removal and deskewing
UniSoft Imaging	Imaging Software Libraries for compression, rotation, and display of image data

1.3 Terms and Abbreviations

ADP	Automated Data Processing
API	Application Program Interface
CAD	Computer-aided Design
CADA	Computer-Assisted Data Acceptance
CALS	Computer-aided Acquisition and Logistic Support
CCITT	Consultative Committee for International Telegraphy and Telephone
CECOM	Communications-Electronics Command
COTS	Commercial-off-the-shelf
CPU	Central Processing Unit
CTC	CALS Technology Center
CTN	CALS Test Network
DIU	Document Image Understanding
DLA	Defense Logistics Agency
DOD	Department of Defense
DSREDS	Digital Storage and Retrieval Engineering Data System
ED	Engineering Data
EDCARS	Engineering Data Computer Assisted Retrieval System
ERR	Engineering Release Report
GB	Gigabyte
GUI	Graphical User Interface
ICR	Intelligent Character Recognition
ID	Identification Data
IMS	Image Management Systems
IPR	In Progress Review
JCALS	Joint Computer-aided Acquisition and Logistic Support
JEDMICS	Joint Engineering Data Management Information Control Systems
LAN	Local Area Network
MB	Megabyte
MHz	Megahertz
MICOM	Missile Command
NIC	Network Interface Card
NFS	Network File System
OASD	Office of the Assistant Secretary of Defense

[1] Equivalent COTS products may be used instead of the products specified (refer to Appendix A).

OCR	Optical Character Recognition
OS	Operating System
PDL	Page Description Language
PM	Program Manager
QA	Quality Assurance
RAM	Random Access Memory
ROI	Region of Interest
ROM	Read Only Memory
SCSI	Small Computer Systems Interface
SU	Software Unit
TCP/IP	Transport Control Protocol/Internet Protocol
VLSI	Very Large Scale Integration
VQA	Visual Quality Assurance

SECTION 2 SYSTEM SUMMARY

2.1 Background

The acquisition, storage, and management of data for the logistic support of weapon systems has long been a challenge for the Department of Defense (DoD). As weapon systems' technological complexity increased, the volume of data expanded and the challenge for acquiring/maintaining quality data became more difficult. The DoD responded to this challenge by authorizing the procurement of Image Management Systems (IMS) beginning for the Army and the Air Force in the early 1980's, and for the Navy in the late 1980's. The DoD is continuing to address the storage and distribution problems of engineering data and is now beginning to replace the 1980's technology-based Service repositories with a common system that can be jointly used throughout the DoD. This system is known as the Joint Engineering Data Management Information Control System (JEDMICS). All data within this system is stored in electronic formats.

As this large volume of data is received at DoD sites, the need for quick and reliable Quality Assurance (QA) rises. Any automated assistance available to existing QA personnel will dramatically aid in reducing the time required to ensure only good quality data is stored within the repositories.

Also, as these existing DoD repositories are upgraded, the need to migrate the existing data from the old systems to the new JEDMICS system must occur. During this transfer of information, it has been found that the contents of the index information is not always consistent. This is a real problem concerning large format drawings which contain multiple sheets. As a result it has become necessary to correct these inconsistencies.

The (CADA) system described herein will address how automated QA techniques will aid QA personnel with their current job and also correct the index information inconsistencies identified to be present during the migration effort.

2.2 Objectives

The CADA system will provide JEDMICS QA personnel with the ability to perform automated QA according to the following two scenarios:

> **PENDING BATCHES** - The user is able to perform image quality and index information verification for a specified batch which is resident in JEDMICS Pending. Using CADA, the user can change index information and QA flags which will subsequently be changed within the given batch in Pending.
>
> **LARGE FORMAT DRAWINGS FROM PERMANENT STORAGE** The user is able to initiate index information corrections for large format drawings which are contained within JEDMICS Permanent.

The objective of CADA, through automation, is to cost-effectively perform 100% QA for batches resident within JEDMICS Pending and correct inconsistencies for large format drawing header information within JEDMICS Permanent

2.3 Existing Methods and Procedures

The existing data acceptance and QA methods and procedures used by the DoD engineering data repositories vary from site to site. A generalized procedure, that encompasses procedures from each of the Tri-Services, is discussed in relationship to the flow of data within a repository site. The IMS repositories currently used are Digital Storage and Retrieval Engineering Data System (DSREDS), Engineering Data Computer Assisted Retrieval System (EDCARS), and JEDMICS.

The general flow of engineering data is shown in Figure 1 and should be referred to during the discussion of a typical procedure that may be followed for the receipt and acceptance of data in aperture card or in magnetic tape formats. As shown in Figure 1, the data are received, inspected, loaded, and temporarily stored in electronic format. Final QA is then performed by visual inspection of the engineering data on high-resolution graphic workstations. Each of these steps are discussed in Sections 2.3.1 through 2.3.4.

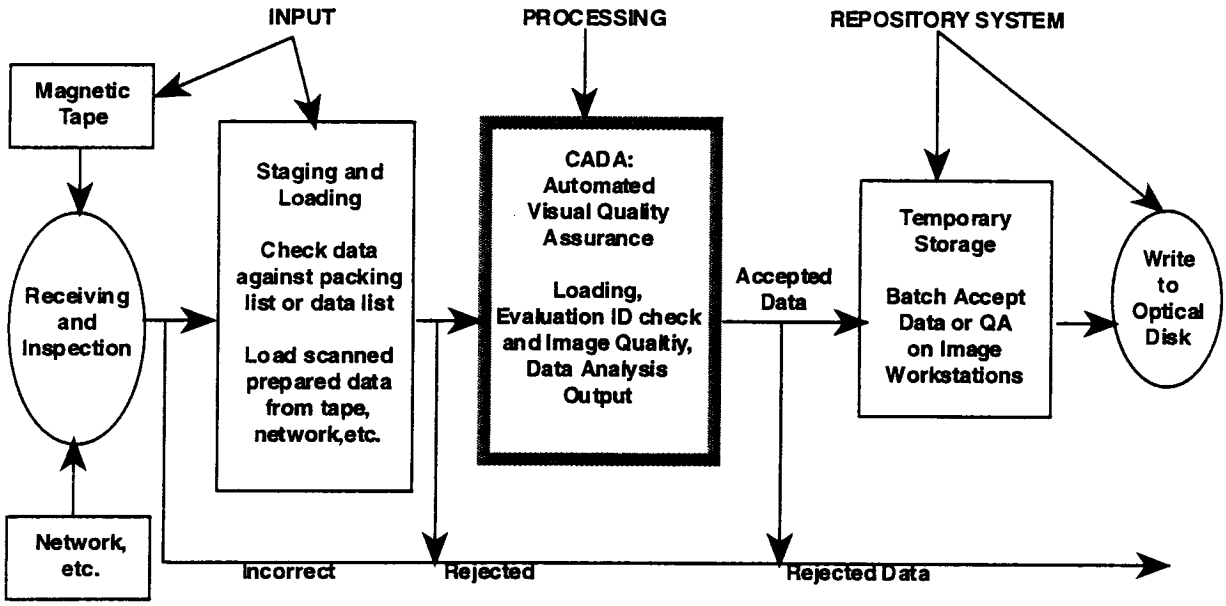

Figure 1. Typical Flow of Data within a Government Repository

2.3.1 Input - Receiving

The Tri-Services' engineering repositories are presently capable of receiving and storing engineering data in hardcopy, microform, and digital raster formats. This scenario will discuss only data on aperture card and magnetic media, since this processing is most common and represents the characteristics for most other format processing. The incoming receiving/inspection personnel verify that the data received matches the shipping/packing list and data list (when available), and that the physical condition of the deliverable is acceptable. The media is then sent to storage, prior to delivery to the engineering data section for staging, before loading into the IMS.

2.3.2 Input - Staging by Engineering Data (ED)

The first evaluation by Engineering Data (ED) prior to entrance into the IMS repository is the recording and preparing of Engineering Release Records (ERR) or equivalent documents. At this point, the procedures vary among the users. Some sites check the accuracy of the punched Hollerith code of the aperture cards and perform visual QA of the image during this step. This is not often done anymore since the IMS software can detect many of these errors during the scanning of the aperture cards. The aperture cards are sorted into batches in preparation for scanning into the system, or, in the case of magnetic tape media, ED personnel may log the number of tapes and prepare an ERR. They then forward the tape with the data list and aperture cads, if required, for loading into the IMS.

2.3.3 Processing - Loading the Data

The scanning operator will enter the batch information into the IMS in preparation for scanning. The aperture cards are then scanned and gross image problems or punched Hollerith errors are detected at the scanning stage. The appropriate cards are rejected. The IMS software builds the index directory either by batch, or in some cases, by card image during scanning. If a batch index is used, then the index by drawing is done at an index workstation during image QA. The IMS software is more often used to build the index by drawing number from the punched Hollerith card during the scanning and loading process. The batch data is temporarily stored on magnetic disk prior to QA. In the case of the magnetic tape, the operator mounts the tape onto the IMS tape drive and loads the data into temporary storage.

2.3.4 Visual Quality Assurance (VQA)

The QA operator will access the batch directory of data to be analyzed from temporary storage and select the first image for QA. At this point, the operator will perform QA in accordance with the local procedures (sampling, 100% QA, etc.). The operator may refer to a hard copy batch data list during the QA and mark the image list for acceptance, rejection, or hold. In some cases, the operator has the scanned batch of aperture cards that can be referred to if questions arise. An image may be placed on "hold" for a number of reasons. The operator may wish to view the image further, have other operators view the image, or obtain additional information before an accept or reject decision can be made. The marked data list, aperture card, or other documentation will show the images accepted from the scanned batch.

This method of Visual QA is time-consuming and open to a subjective decision, depending on the background and experience of the QA operator. Obviously, manual VQA can be costly, especially if 100% visual QA is required. The results are totally dependent on the operators' QA knowledge and application of that knowledge.

2.4 Proposed Methods and Procedures for the Application of CADA Techniques to JEDMICS

This system is proposed as a means to assist existing QA procedures and improve the quality of index header information contained within the JEDMICS Tri-Services' repositories. It is intended to add the processing, described in Sections 2.4.1 through 2.4.2.3 on a separate hardware platform and

develop an interface from this platform to the JEDMICS Application Program Interface (API) server. This will allow for the introduction of this processing with minimal impact on the current JEDMICS flow of operations.

This section addresses, at the highest level, the functions to be performed. Figure 2 shows the proposed data flow.

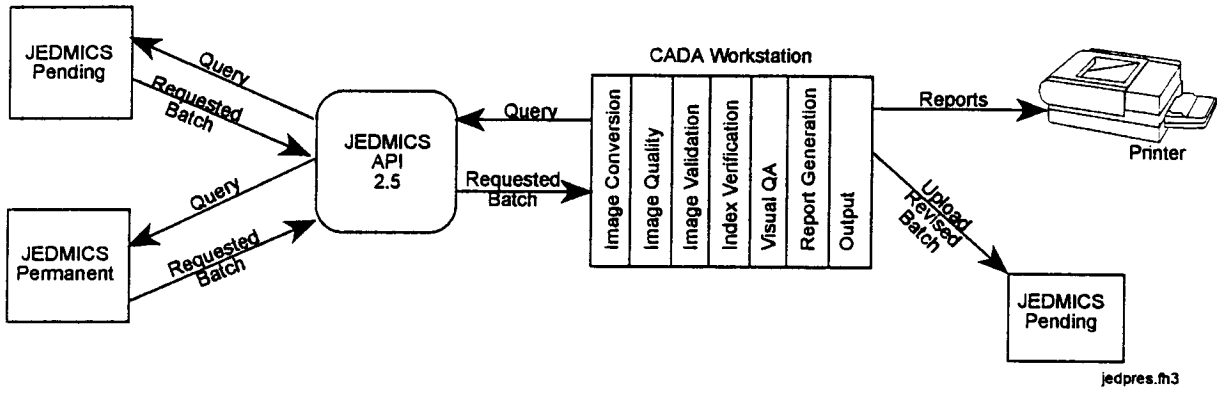

Figure 2 - JEDMICS/CADA Data Flow

The JEDMICS/CADA workstation will request batches of images from pending or permanent storage via the JEDMICS API. The JEDMICS/CADA user is responsible for initiating the request. Once received at the JEDMICS/CADA workstation, the batch will undergo the appropriate processing based upon users request.

If the user requested a batch from Pending, then the batch may undergo the entire CADA processing. This consists of image conversion, image quality, image validation, index verification, visual QA, report generation, and batch output The extent of the automated image quality and index verification may be further restricted by the CADA user.

If the user requested a batch from Permanent then the batch will only undergo the image conversion, a subset of the index verification (to check for inconsistencies with large format drawing), visual QA, report generation, and batch output.

Once CADA processing is completed, CADA will generate a new batch containing the changed information which will be submitted to Pending. This will include the setting of QA flags and any necessary changed index information.

During the formation of this design it was thought that the upload operation may be better if it interfaced directly with JEDMICS Permanent. This may be a consideration in the future due to the processing time it would eliminate at Pending.

This direct method has been put on hold for the following reasons:

- maximum use of the JEDMICS API can be made,

- the interface to the JEDMICS Permanent is not fully documented, and
- upload via Pending allows operators to perform a final check prior to re-insertion to JEDMICS Permanent.

2.4.1 Summary of Improvements

This system will provide a number of improvements over how the same processing would be accomplished using an equivalent manual effort.

- Functional Improvements

The capabilities of JEDMICS/CADA provide the QA operator with tools that automate the image QA of engineering drawings in an unattended mode to maximize the QA operators' ability to efficiently and accurately arrive at image quality decisions. The operators analyze the JEDMICS/CADA results and apply their expertise to make the final QA decisions.

JEDMICS/CADA also provides the capability to automate the analysis of index information within a drawing and compare it with the corresponding index information used to store the image. Again, the operator has the opportunity to analyze the JEDMICS/CADA results and override them if he or she chooses to do so. These index checks will be expanded to check for inconsistencies in index information for large format drawing which reside in Permanent.

The index information checks will have a false accept rate of zero percent.

Other functional improvements include a Graphical User Interface (GUI) that provides the operator with flexibility in retrieving, displaying, and analyzing the JEDMICS/CADA results or any engineering drawing in the batch being analyzed. A key function is that the CADA system can be used in a stand-alone mode without impacting the Image Management System repositories' existing QA functions.

- Improvements of Degree: The engineering data section's performance can be improved by providing the stand-alone capability for unattended, automated QA of engineering drawing data, thereby allowing other work to be performed concurrently. JEDMICS/CADA can be used with the expertise of the QA operator to perform full acceptance of the data and forward the accepted data and appropriate batch and QA flags to JEDMICS Pending.

  Timeliness: Presently, CADA operates in a stand-alone mode and does not impact the response time of the repository system in any way. The overall response time for the QA of batches of data will be improved by the addition of the CADA system at the Government JEDMICS site for pre-acceptance of batches or for the correction of index information inconsistencies.

2.4.2 Summary of Impacts

Implementation of the JEDMICS/CADA system within the procurement and acceptance of engineering drawing data impacts the organization, the operational procedures, and the user.

The design of CADA is based upon the QA functions performed in the existing systems. Since CADA is a stand-alone system, its impact on users is confined to the QA process of reviewing data intended for storage on a repository system. The application of the CADA tools by QA personnel should result in a cost-effective acceptance of engineering drawing raster data. JEDMICS/CADA will impact QA operators in the following areas:

- uniformity of QA process,
- efficiency of QA process, and
- training to effectively use JEDMICS/CADA.

The costs of implementing JEDMICS/CADA are spread across hardware equipment, software (third party licenses), user training, and limited integration efforts. Various price/performance options are available when acquiring the hardware necessary to run JEDMICS/CADA. The software costs of JEDMICS/CADA are confined to the licensing and use of third party products needed to use JEDMICS/CADA (refer to Appendix A). User training has been straightforward in the past since JEDMICS/CADA has a user-friendly GUI that requires minimal effort for the user to operate. JEDMICS/CADA provides various ways of accomplishing basic tasks and can be easily configured to operate in the fashion that best suits the user.

2.4.2.1 User Organizational Impacts

The application of CADA in the procurement of weapon systems data from the contractor could result in modification of the QA acceptance of data at the contractors site. The manual effort of Government inspectors would be reduced since the contractors' data could be analyzed by JEDMICS/CADA and the inspector could analyze the output report to make an accept/reject decision for pre-acceptance. If the contract required the contractor to use JEDMICS/CADA to perform 100% QA analysis of the data, the task of on-site QA inspectors would be greatly simplified. The obvious result would be improved quality of data for the Government. Organizational impacts at the Government repository will mostly be improved QA throughput, and more efficient use of the QA operators time in that 100% visual QA will not be required of the operator. The engineering data section methods of operation will not be greatly impacted since JEDMICS/CADA is not presently integrated within the repository system.

There are no user organizational impacts as a result of JEDMICS/CADA's use in the QA process. There would not be a modification or elimination of positional responsibilities. QA operators will need to undergo training to use the JEDMICS/CADA software. Training sessions will specify the strengths and differences JEDMICS/CADA imposes on the typical QA process. A system/security administrator would be required to assist in configuring the JEDMICS/CADA workstation. The administrator should be responsible for controlling who has access to the system, and training users on basic security precautions when using a multi-user computer operation system.

2.4.2.2 User Operational Impacts

JEDMICS/CADA, as a stand-alone system, will not have a great impact on the operation of the repository organization, since direct integration into the JEDMICS repository will not be part of this effort. The current operational procedures within the engineering data, or acceptance section of the repository, will be minor in that JEDMICS/CADA will retrieve and analyze the data with little operator input. QA personnel will analyze the automated results and direct JEDMICS/CADA to submit the changed information to Pending.

This procedure would be an addition to the existing procedures and would reduce the visual QA presently required within the repository. The required aspects of data acceptance (legibility, reproducibility, and accuracy of data) would not change. The use of JEDMICS/CADA impacts where and the way QA is performed.

JEDMICS/CADA would be used as a "first-stop" QA station where data would be reviewed before being input into the repository system. Currently, incoming data are loaded into a temporary location on the repository system and reviewed by a QA operator at an image workstation. Accepted documents are written to the optical disk of the repository system and a rejected document's report is sent back to the originating site. Using JEDMICS/CADA, incoming data would be reviewed at the JEDMICS/CADA workstation, then the QA operator would forward the accepted documents to Pending for loading to the repository system. A rejected document's report produced by JEDMICS/CADA (notes and reject reasons are added during QA) would be sent back to the originating source.

CADA will also be used to correct index inconsistencies found within Permanent for large format drawings. These inconsistencies will be detected and corrected by JEDMICS/CADA. QA personnel will then direct JEDMICS/CADA to submit any changes to Pending.

JEDMICS/CADA addresses the underlying need of providing a foundation for uniform and consistent acceptance criteria for QA procedures. JEDMICS/CADA helps to factor out the human subjectivity component of QA analysis. Subjectivity exists among different users and also for a single user across a time span. Common causes are fatigue, human errors, and outside influences (such as the quality of previously viewed images or display device limitations). JEDMICS/CADA provides a consistent starting point for reviewing images by judging all images in a batch with a consistent criteria. By providing an automated system which performs QA checks without intermediate human intervention, QA throughput is increased.

Cost-effective 100% QA of engineering drawing raster data can be performed since QA operators are armed with JEDMICS/CADA's automated evaluation information about each image in a batch and can develop reports interactively while reviewing documents. As users become more confident and familiar with JEDMICS/CADA, QA efficiency can be increased by viewing only the most relevant images in a drawing (e.g., first, second, and last sheets in a document). In addition, since operators have the ability to override CADA decisions, final accept/reject control is always retained by the operator.

JEDMICS/CADA will also assist the users in compiling reports about incoming data. At any time during the QA process the user can generate a report that lists information about each image in a

batch including its index information, user generated on-line comments, and a final reject/accept decision.

2.4.2.3 User Development Impacts

The user of JEDMICS/CADA requires familiarity in the use of a keyboard and mouse. Limited training will be required for the user to understand the documentation provided and the GUI. Training sessions will focus on how current QA practices map over to the QA of documents using JEDMICS/CADA. The users will initially enter data to identify the batch and the contract information necessary to track the data and then return after the JEDMICS/CADA analysis to review the screen or hardcopy results. The operator may then choose to sample the JEDMICS/CADA rejected image or index information drawing images for override or any other disposition he/she may choose. The QA operator would forward the accepted documents to Pending.

2.5 Assumptions and Constraints

2.5.1 Assumptions

The following is the list of assumptions which dictate the completion of the CADA/JEDMICS interface:

- A JEDMICS API server will exist at the site where the JEDMICS/CADA platform is to be installed.

- A separate hardware platform (outlined in Section 5 of this document) is available for the client/CADA software.

PENDING BATCH OPERATIONS

- The user has knowledge of the batch number (assigned by the Pending data base) which is to be processed.

- The user has access to the JEDMICS/CADA workstation and authorization to perform QA against the Pending data base.

- The user has the responsibility to release the JEDMICS/CADA results to pending storage.

- The image quality and index information checks which are performed and reported to pending are only those addressed in Section 4 of this document.

PERMANENT STORAGE OPERATIONS

- The index information which will be updated for large format drawings is limited to the drawing size.

- Large format drawings are correctly indexed within JEDMICS.

- Drawing revision extraction is limited to C, D, E J, G, and H size (multiframe) drawings.

- Drawing revision extraction will be for all revisions of each large format drawing identified.

2.5.2 Constraints

The following is the list of constraints which dictate the completion of the JEDMICS/CADA interface:

- The download operations will be directed to the server through the JEDMICS API.

- No updates will be performed for image files.

- The upload operations will be directed to the server through the JEDMICS API.

PENDING BATCH OPERATIONS

- JEDMICS/CADA QA operations will be performed against only one batch at a time.

PERMANENT STORAGE OPERATIONS

- The upload operation will utilize Pending and will therefore, re-insert only the index header information.

- The download will retrieve the image as well as the associated header information for each requested engineering drawing revision.

SECTION 3 DETAILED CHARACTERISTICS

This section provides a detailed description of the functions to be performed and the performance requirements of the proposed software. The characteristics are specified in accordance with MIL-STD-1835A (format).

3.1 Specific Performance Requirements

The following requirements for the proposed system encompass four areas: batch retrieval/image query building, image analysis, index information verify/update, and batch submission.

- Unattended operation for automated evaluation of image legibility and reproducibility
- Ability to process JEDMICS C4 format images
- Low false accept error rates for image analysis
- Low false reject error rates for image analysis
- Unattended operation for automated evaluation of validity of index information
- Provide capability to process engineering drawings without borders
- Low false accept error rates for index information analysis
- Low false reject error rates for index information analysis
- Map CADA errors to JEDMICS errors
- Provide user-friendly GUI to perform routine tasks
- Provide easy configuration of the application
- Provide alphanumeric sorting of documents
- Provide viewing of images at various zoom levels with flexible panning options
- Allow QA personnel to override automated decisions
- Provide capability of entering notes about images and to print reports
- Allow viewing and modification of index information and Hollerith data
- Allow interactive report generation and printing of final accept/reject decisions

- Provide ways to sort viewing of images based upon various criteria
- Use a high resolution monitor for viewing images
- Provide flexible design that allows incorporation of new features and software into JEDMICS/CADA

3.1.1 Accuracy and Validity

The accuracy and validity concerns of the proposed software system are focused on the automated evaluation results generated by JEDMICS/CADA. The following goals have been targeted.

- A false image query building error rate less than one percent.
- A false image quality accept rate less than one percent.
- A false image quality reject rate less than ten percent.
- A false index information verification accept rate of .01 percent.
- A false index information verification reject rate of less then 40 percent.
- A false index information update error rate of zero percent.
- A false index information re-insertion error rate of zero percent.

This system will address the need of ensuring good quality image files and the need to correct large image format index information. Due to automation, it will help to reduce the time factor associated with such jobs.

3.1.2 Timing

The proposed system will have the following requirements for throughput.

- Provide unattended automated batch evaluation that can process about 1000 images overnight.
- Provide fast viewing of images: no more than 12 seconds for full scale view of E size images under regular system load.
- Provide fast generation of evaluation and data list reports: no more than ten seconds for a 100 image batch report under regular system load.

Performance of the JEDMICS/CADA software will be dependent on several factors. Most importantly is the processor speed and amount of Random Access Memory (RAM) installed in the workstation. The Central Processing Unit (CPU) and memory-intensive operations of the automated evaluation and decompression of images for viewing will be greatly affected by these

Image Conversion: This function is responsible for the conversion of the JEDMICS C4 formatted image files to TIFF format. All errors will be reported accordingly. This conversion will be done at the time the batch is loaded into JEDMICS/CADA. Since there will be no updates for image files the reverse conversion is not needed.

Evaluate Index Information: This function is responsible for performing automated evaluation of all index information. The accuracy of the index information will be compared against the same information which is present on the image face. The user will only have to initiate the evaluation. Once started, the evaluation will run unattended.

This function is available to be run only against batches downloaded from the Pending data base.

Evaluate Large Format Drawing (LFD) Index: This function is responsible for identifying a large format drawing and checking the drawing size field contained within the index information. If this information is not correct it will automatically be corrected.

This function is available to be run against batches downloaded from Pending as well as Permanent.

Evaluate Image Quality: This function is responsible for checking the image quality for reproducibility. Any problems encountered with the image will be reported to the user.

This function is available to be run only against batches downloaded from the Pending data base.

3.3 Inputs and Outputs

The JEDMICS/CADA system will accept the input of JEDMICS C4 image files and the corresponding index information for each of these files. All inputs will be addressed within JEDMICS/CADA as a batch. If the batch is received from JEDMICS Pending, JEDMICS/CADA will retain the batch number for consistency but will assign its own batch number for internal processing. If the batch is received from Permanent, JEDMICS/CADA will assign a unique batch number for processing within JEDMICS/CADA.

The output from the system will be a set consisting of the index information which is to be changed. There will be specific changes related to the batch and also to each individual image within the batch. The following is a complete list of the parameters which are candidates for change:

BATCH Object
JMX_batchId*
JMX_badHollerithCount

JMX_blurredCount
JMX_deletedCount
JMX_flaggedCount
JMX_indexedCount
JMX_migratedCount*
JMX_poorCompCount
JMX_qaFailureCount
JMX_tooDarkCount
JMX_tooLightCount
JMX_unsureCount
JMX_batchCount*
JMX_dateBatchCreated*

DRAWING Attributes
JMX_documentType
JMX_drawingCount
JMX_drawingNumber
JMX_drawingRevDate
JMX_drawingRevision
JMX_drawingSize*
JMX_drawingTitle
JMX_frameCount
JMX_frameNumber
JMX_numberOfFrames
JMX_numberOfSheets
JMX_qaFlags
JMX_securityLevel
JMX_sheetCount
JMX_sheetNumber
JMX_sheetRevision
JMX_pendingStatus

* These fields are the only fields which will be set for a batch originating from permanent storage.

For batches which originated from the Pending data base, the above information will be updated for the original batch. For batches which originated from Permanent, the above information will be supplied in a newly created batch within Pending. The changed index information will be added as a new drawing. The old information will be tagged for deletion.

3.4 Data Base/Data Bank Characteristics

JEDMICS/CADA will use various flat file indexes/data banks to store information about loaded batches and images. A batch information index will maintain information relevant to the loaded batches. This information will include a batch's load time, size, declaration file contents, evaluation results, and identification information specified by the user. Another index will maintain information relevant to each image file present in a batch. This information will include an image's

index information, evaluation results, user generated comments, and raster data information. Other indexes will store additional information about batches such as alphanumeric sort orders, etc.

3.5 Failure Contingencies

The following describes the suggested failure contingencies for the proposed system.

Data Backup - No special data backup procedures should be required by JEDMICS/CADA. Backups of JEDMICS/CADA's temporary batch storage and log file directories could be performed by the system administrator during regularly scheduled backups. The system administrator should store JEDMICS/CADA's installation tapes or floppy discs to reinstall the software, if necessary.

Fall Back - JEDMICS/CADA will need to display various messages throughout the application when errors occur. Most of the messages should be self-explanatory and non-fatal. The user should be able to deal with this situation which usually involves bad data entry on the user's part. The user should be able to refer to the JEDMICS/CADA End User's Manual for information on how to deal with basic JEDMICS/CADA error messages. Any errors that are serious in nature should be specified and the system administrator should be able to handle this situation. The system administrator should be able to refer to the JEDMICS/CADA Maintenance and Computer Operator's manuals for instructions on how to deal with fatal situations.

Degraded Modes of Operation - There will be no degraded modes of operation.

SECTION 4 DESIGN CONSIDERATIONS

4.1 System Description

JEDMICS/CADA is described as a tool that will provide for the automated QA of engineering drawing raster data. It will act as an unattended, objective, and uniform quality method of evaluating raster data for legibility and reproducibility in addition to providing automated validation of the index information contained within the body of the drawing with equivalent index information contained in the file's header.

JEDMICS/CADA will allow a user to retrieve a batch of data, perform automated evaluation of images, view/print the images, override the automated results if necessary, generate reports, and resubmit the results to JEDMICS Pending. JEDMICS/CADA will operate in a stand-alone environment on a Sun SPARCstation that has access (locally or via a network) to a printer. The operator initiates the loading and automated processing of the image data, analyzes the results, and views the images to make final decisions for the batch. Once an evaluation and review has been completed, the user can output the changed index information to JEDMICS Pending. JEDMICS/CADA provides a user-friendly GUI that requires minimal training for the user to operate.

The use of JEDMICS/CADA within the existing JEDMICS environment will be developed as a stand-alone system, with an interface to the existing JEDMICS repositories that is confined to the extraction of batches from Pending or large format drawings from Permanent and the inputting of changed index information to the Pending data base.

4.2 System Functions

4.2.1 Batch Retrieval

JEDMICS/CADA will use the existing JEDMICS API to retrieve batches from either Pending or Permanent, based on the operations to be performed.

When retrieving a batch from Pending, the user is expected to know the batch number (given by the Pending data base) of the batch he/she wishes to QA using JEDMICS/CADA. This batch of images must have already been loaded into Pending using the normal methods currently being used by JEDMICS QA personnel.

The following API calls will be used when retrieving a batch from pending storage:

JMXCreateSession This is used to setup the environment for log in
JMXOpenSession This is used to log into the server
JMXCreateBatch This is used to allocate resources for the batch
JMXSetCriteria This is used to set the batch number
JMXQueryBatch This is used to check the existence of the requested batch
JMXQueryPendingStorage Once the batch has been identified, this function call will be called to retrieve all images index information and images within the requested batch
JMXRetrieveImage This is used to retrieve individual image files
JMXCloseQuery
JMXCloseSession
JMXDestroySession

The steps involved for the downloading the images into CADA are described below. These steps will be performed by the JEDMICS client process running on the Sun workstation.

1. A session is established in order to communicate with the JEDMICS server. The API call used for this purpose is JMXCreateSession.

2. A batch object is established via the JMXCreateBatch API call.

3. A drawing handle is allocated, that would store the information regarding the drawings to be queried. The API call used for this purpose is JMXCreateDrawing.

4. The query values for obtaining a specific batch are input by setting the appropriate parameters.

For example: In order to retrieve all the drawing with the drawing size of J, JMX_drawingSize parameter is set. The API call used for this purpose is JMXSetValue. The call would then be:

JMXSetCriteria(drawing, JMX_batchId, (JMXPointer)"1634");

Where drawing is the handle returned in step 3, JMX_batchId is the handle returned in Step 2, and the criteria for the search is set to find batch with ID 1634.

4. A query is then submitted to JEDMICS pending storage. The API call used is JMXQueryPendingStorage.

   The index information relating to all the drawings that satisfy the query will then be downloaded.

5. This index list is then parsed and each drawing image is sequentially retrieved. This drawing image is now available for processing, along with its index information.

   The API calls JMXGetNextDrawing, JMXGetNextSheet and JMXGetNextFrame are used while parsing the list of retrieved drawings. The API call used to extract the actual image from the data base is JMXRetrieveImage.

6. After all the drawings are processed by CADA, the drawing structure is deallocated using the API call JMXDrawingCloseQuery.

7. The JEDMICS session is then closed by the client using the API call JMXCloseSession.

8. The drawing object is then destroyed using the API call JMXDestroyDrawing.

9. The session object is destroyed using the API call JMXDestroySession.

When retrieving a batch from permanent storage the user is expected to enter certain criteria specifying the scope of large format images to be tested. The options which will be allowed for specifying the criteria are the following:

- All This will search the entire permanent storage for large format drawing candidates,

- Specific* The user will then enter more specific criteria to limit the search;

  - CAGE Code
  - Drawing Size, and

- By Date* The user will then enter a range of dates which specifies a load date from which to start the search.

*** A combination of Specific and By Date can also be specified.**

JEDMICS/CADA will then formulate the appropriate queries to retrieve the requested information.

The following API calls will be used when retrieving a batch from Permanent :

JMXCreateSession	This is used to setup the environment for log in
JMXOpenSession	This is used to log into the server
JMXCreateBatch	This is used to allocate resources for the batch
JMXSetCriteria	This is used to set the batch number
JMXCreateDrawing	This is used to allocate resources for drawings to be retrieved
JMXQueryPermStorage	Once the batch has been identified, this function call will be called to retrieve all images index information and images within the requested batch
JMXRetrieveImage	This is used to retrieve an individual image
JMXCloseQuery	
JMXCloseSession	
JMXDestroySession	

The steps involved in downloading the images from JEDMICS Permanent into JEDMICS/CADA are described below. These steps will be performed by the JEDMICS client process running on the Sun workstation.

1. A session is established in order to communicate with the JEDMICS server. The API call used for this purpose is JMXCreateSession.

2. A drawing handle is allocated that would store the information regarding the drawings to be queried. The API call used for this purpose is JMXCreateDrawing.

3. Appropriate query values are input by setting the appropriate parameters.

   For example: In order to retrieve all the drawing with the drawing size of J, JMX_drawingSize parameter is set. The API call used for this purpose is JMXSetValue. The call would then be:

   JMXSetCriteria(drawing, JMX_drawingSize, (JMXPointer)'J');

   Where drawing is the handle returned in step 2.

4. A query is then submitted to the JEDMICS permanent storage. The API call used is JMXQueryPermStorage.

   The index information relating to all the drawings that satisfy the query will then be downloaded.

5. This index list is then parsed and each drawing image is sequentially retrieved. This drawing image is now available for processing, along with its index information.

The API calls JMXGetNextDrawing, JMXGetNextSheet and JMXGetNextFrame are used while parsing the list of retrieved drawings. The API call used to extract the actual image from the data base is JMXRetrieveImage.

6. After all the drawings are processed by CADA, the drawing structure is deallocated using the API call JMXDrawingCloseQuery.

7. The JEDMICS session is then closed by the client using the API call JMXCloseSession.

8. The drawing object is then destroyed using the API call JMXDestroyDrawing.

9. The session object is destroyed using the API call JMXDestroySession.

For batch retrieval from permanent storage, JEDMICS/CADA will need to perform additional management of the images. This involves splitting the user request into smaller batches if the amount of data is too large for handling by the JEDMICS/CADA workstation. In this case, the integrity of the user request is limited to the time/date when the request was submitted.

4.2.2 Batch Submittal

JEDMICS/CADA will use the existing JEDMICS API to submit batches to the Pending data base. The manner in which this is done and the information which JEDMICS/CADA will change is based on the original location of the batch (Pending or Permanent).

Due to restrictions of the API, Pending batch information cannot be modified. Therefore, when processing Pending batches, JEDMICS/CADA will create a new batch with all of the original information, plus the changes approved from automated QA at the JEDMICS/CADA workstation. When processing Large Format Drawings, JEDMICS/CADA will create a new batch which will consist of the changed index information which will be added to the permanent storage along with the original index information which will be tagged as deleted from Permanent.

As part of this output to the Pending data base, JEDMICS/CADA will set or update all appropriate QA flags needed by JEDMICS. These flags are listed in Section 3.3.

JEDMICS/CADA will also allow the QA personnel to output reports which summarize the results of the automated QA.

When submitting a batch to the Pending data base, JEDMICS/CADA will create a new batch and supply the necessary information for the batch. This is due to the current operations allowed while using the API. The following API calls will be used when submitting a batch to Pending:

JMXCreateSession	This is used to setup the environment for log in
JMXOpenSession	This is used to log into the server
JMXCreateBatch	This is used to allocate resources for the batch
JMXOpenBatch	This is used to create a new batch
JMXBatchInsertFile	This is used to put image files into a specified batch

JMXCloseQuery
JMXCloseSession
JMXDestroySession

The steps involved in submitting the images from JEDMICS/CADA to JEDMICS Pending are described below. These steps will be performed by the JEDMICS client process running on the Sun workstation.

1. A session is established in order to communicate with the JEDMICS server. The API call used for this purpose is JMXCreateSession.

2. A batch object is created that would store the information regarding the new batch. The API call used for this purpose is JMXCreateBatch.

3. A new batch is then created within JEDMICS pending storage. The API call used for this purpose is JMXOpenBatch.

4. A drawing handle is allocated, that would store the information regarding the drawings to be inserted. The API call used for this purpose is JMXCreateDrawing.

5. Appropriate values are set in the drawing handle for the image to be submitted. For example,

   JMXSetValue(drawing, JMX_drawingSize, JMXPointer(image size));

   where image size is retrieved from the corresponding field in the internal CADA data structure for the image to be inserted.

6. The image is then inserted into the currently open batch using the API call JMXBatchInsertFile.

7. The Steps 5 and 6 are then successively repeated for all the drawing images to be inserted into the batch.

8. The batch is then closed using the JMXCloseBatch call.

9. The JEDMICS session is then closed using the JMXCloseSession API call.

10. The drawing object is then destroyed using the API call JMXDestroyDrawing.

11. The batch object is destroyed using the API call JMXDestroyBatch.

12. The session object is destroyed using the API call JMXDestroySession.

4.2.3 Input/Output

The input of batches downloaded from JEDMICS occurs strictly for allocating and populating internal data structures required by JEDMICS/CADA. Upon input, the index information will be

appropriately stored and the batch ID will be established. This ID will be retained from pending storage or assigned if the batch is received from Permanent.

The output of batches which have been processed through JEDMICS/CADA involves establishing the appropriate set to be submitted to Pending. The output set will consist of only the index information, as the image files are not changed by JEDMICS/CADA. This index information will be appropriately set based on the operations performed within JEDMICS/CADA.

The output will also involve the mapping of JEDMICS/CADA results to QAFlags understood by JEDMICS. The following table shows this mapping.

Image Quality Errors

JEDMICS/CADA Error	JEDMICS/CADA Error Description	JEDMICS QA Flag
DECOMP_ERR	Detected when JEDMICS/CADA has had a problem decompressing the image.	JMX_POOR_COMP
HIGH_BLK_ORPHS	This is the ratio of black run lengths of one in non-vertical mode to the total number of white pixels. It approximately indicates the black orphan pixels within the image.	JMX_TOO_DARK
HIGH_WHT_ORPHS	This is the ratio of white run lengths of one in any mode to the total number of black pixels. It approximately indicates the white orphans pixels within the image.	JMX_TOO_LIGHT
HIGH_FF	This technique takes the fill factor and compression ratio of the compressed image into consideration. The assumption is that a noisy image (high fill factor) is compressed to a lesser degree (low compression ratio).	JMX_TOO_DARK
LOW_FF	This technique takes the fill factor and compression ratio of the compressed image into consideration. The assumption is that a sparse image (low fill factor) is highly compressed (high compression ratio).	JMS_TOO_LIGHT
HIGH_RLRAT	Run length analysis is a means of detecting noise data in compressed image data. A high ratio indicates a noisy image. A low ratio indicates a noise-free image.	JMX_TOO_DARK

JEDMICS/CADA Error	JEDMICS/CADA Error Description	JEDMICS QA Flag
HIGH_PEAK_TILE	Technique used to detect images that are faded in small areas. Image is divided into 256 x 256 pixel tiles and the noise level is detected for each tile.	JMX_TOO_LIGHT

If the batch was originally received from Pending and image quality and index verification were performed then a new batch would be generated which would consist of all updated information from within JEDMICS/CADA as well as all unchanged information retained by JEDMICS/CADA.

If the batch was originally received from permanent storage then the updated index information would be submitted as new and the index information it is correcting would be tagged to be deleted.

4.2.4 Image Conversion

The image conversion is performed upon the input of a batch within JEDMICS/CADA. Since JEDMICS/CADA performs all image processing on TIFF formatted files the JEDMICS C4 image files must be converted to TIFF. All errors will be reported accordingly. Since there will be no updates for image files the reverse conversion is not needed.

4.2.5 Evaluate Index Information

The evaluation of the index information will perform the tasks of matching the index information received from Pending or Permanent with the corresponding ID contained within the raster image. It will locate the index information in the image, uses Intelligent Character Recognition (ICR) software to recognize the data, and matches the recognized data with the index information associated with the image file. The index fields that will be verified are: drawing type, drawing number, CAGE code, sheet number, revision letter, and size letter.

In addition to this checking, all images which are candidates for large format drawing (LFD) index verification will be checked accordingly. This checking is discussed in Section 4.2.6.

4.2.6 Evaluate Large Format Drawing (LFD) Index

Images will be considered candidates for large format drawing processing if they are of size C, D, E, J, G or H. The physical size of the image will be cross-verified with the size contained in the corresponding index information. The large format drawings are split across multiple frames. These multiple frames have the identical index information (as per their six ID fields mentioned in the last section), making them candidates for large format drawings. These multiple frames will then be evaluated for additional characteristics like continuation title block and a determination made as to its large format size. The size field in the index information will then be updated, if necessary.

4.2.7 Evaluate Image Quality

The first step performed as part of the image quality evaluation is image validation. This effort involves performing preliminary Document Image Understanding (DIU) on the image file. As a minimum, this consists of deskewing the image and performing a physical size check, drawing border location, and image classification. For image classification, the document face will be identified to be either an engineering drawing or associated list type.

The next step is to perform the image quality analysis. This is handled by various pixel-level algorithms such as approximate black orphan analysis, approximate white orphan analysis, peak tile noise, run length ratio, fill factor, etc. The results of the algorithms will be used either independently or combined by using a decision tree to arrive at a final image quality ACCEPT or REJECT decision.

4.2.8 Visual QA

The visual QA workspace will allow the user, as a minimum, to display individual images from a batch, override the JEDMICS/CADA evaluation results, enter notes about an image, and modify the image's index and Hollerith information. The workspace will have an image canvas which provides various ways to view and pan around the contents of the image. The visual QA workspace will also display the information about the current image being viewed. This information consists of the image number in the batch, evaluation status, override status, button to select for performing an override, and rejection reason or comment window. The workspace will also display the index or Hollerith (product data) information (drawing number, sheet number, weapons system code, rights, etc.) in editable text fields. If the index information in the image is found to be inconsistent with the displayed header data, by typing in the correct data, the user will be able to correct that index information for the image.

4.2.9 Report Generation

The user will have the capability to display or print evaluation status reports or batch data lists. The Evaluation Status Report will be a snapshot of the active batch. The report will specify the overall statistics of a batch as well as under what conditions the batch was evaluated. The Evaluation Status report will also list information about each image in a batch including its index information, final reject/accept decision, and user generated on-line comment or header modification, if present. The Batch Data List Report contains general information on a batch consisting of the Batch Summary and a Batch Image List Summary.

4.3 Flexibility

The design of JEDMICS/CADA is centered around several distinct modules, the majority of which provide basic input of data/parameters and output of results. Because of this, a flexible control flow of data will be implemented that allows incorporation of new features and software into JEDMICS/CADA without a major effect on overall operation. Because of the modular design, replacement or new COTS products can be incorporated into the software. This allows developers to enhance JEDMICS/CADA's existing capabilities with better COTS products or to add new features as needed.

JEDMICS/CADA has been developed on a Sun Unix platform using C language and X windows/motif for the GUI, the software is very portable to other Unix platforms (the NestorReader and ScanFix COTS products specified in Section 1.2.4 or other products, if used, would need to be available on the desired port platforms; the UniSoft COTS product is available as source code by license and can be easily ported to the desired port platforms). Due to the capabilities of X windows, JEDMICS/CADA can be used remotely from any PC/Unix workstation/terminal-based X server on a network.

JEDMICS/CADA has been designed to allow the user flexibility in customizing the operation of the software. Several functions/parameters can be turned on/off or set to default values when applicable. These options control various aspects of JEDMICS/CADA's operation such as where data and log files are stored, automated evaluation parameters, and device locations/configurations. At least, configuration windows will be provided to handle evaluation specific options and all other configurable features.

4.4 System Data

JEDMICS/CADA makes extensive use of the index information for storing, retrieving, and manipulating the image files. JEDMICS/CADA also stores extensive information for each image within the batch as well as general information on the entire batch. Specific image information includes automated evaluation results and QA personnel overrides. Information regarding the entire batch includes statistics on automated evaluation results regarding poor and good quality images and index information as well as QA personnel operations performed.

Internal file indexes which are used to maintain this information about loaded batches and images are not directly accessible by the user.

SECTION 5 ENVIRONMENT

5.1 Equipment Environment

This section provides a description of the equipment capabilities required for the operation of the JEDMICS/CADA software.

5.1.1 CPU

A Sun SPARCstation with at least 20 megabytes (MBs) of free hard disk space for application related files. A Sun Classic SPARCstation compatible with a 50 Megahertz (MHz) SPARC CPU or better is recommended. A Sun compatible keyboard and mouse is also needed to use the JEDMICS/CADA user interface.

5.1.2 RAM and Hard Disk

The JEDMICS/CADA system requires at least 16 MBs of RAM. JEDMICS/CADA will perform best with 32 MBs or more of RAM. Additional hard disk space is needed to load image data. This space needed is dependent on the size of the batches to be loaded. At least 80 MBs of space is recommended for temporary image data storage.

5.1.3 Display

A high resolution monitor and video card capable of an X windows display. A monitor and video card capable of 1600 x 1280 resolution is recommended.

5.1.4 Data Input/Output

All data input and output will be performed across a Local Area Network (LAN).

5.1.5 Hardcopy Output

A PostScript printer to print reports, batch data lists, and raster images. JEDMICS/CADA can output raster files at a resolution of up to 600 dpi on paper ranging from letter to tabloid size. A printer capable of this resolution is recommended to print out larger size images (D size and up) that need to be scaled down to fit on the paper.

5.2 Support Software Environment

The system's control program coordinates the use of system resources and executes all Input/Output (I/O) instruction. It also handles exceptional conditions, and supervises scheduling and execution of multiple programs. The Sun Operating System (OS) software and ancillary software products are supplied and/or licensed by the manufacturer or software vendor.

5.2.1 Operating System (OS)

JEDMICS/CADA requires Sun OS version 4.1.x (Solaris version 1.x).

5.2.2 Graphical User Interface (GUI)

JEDMICS/CADA requires OpenWindows v3.x or Motif v1.x for the GUI.

5.2.3 Miscellaneous System Setup

At least 1.5 times more swap space than the available RAM in the system (e.g., with 16-mb of RAM, there should be at least 24-mb of swap space). The disk volume that has been designated as the printer spool area should have at least 10 MB of free hard disk space.

5.3 Communications Requirements

JEDMICS/CADA must interface with the JEDMICS server. In order to achieve this communication, the API client software must be resident on the JEDMICS/CADA workstation. There must be a TCP/IP network link between the JEDMICS/CADA workstation and the JEDMICS server.

5.3.1 Graphic Overview

JEDMICS/CADA can take advantage of Transport Control Protocol/Internet Protocol (TCP/IP) network communications by using X windows and Network File System (NFS) protocols. Since CADA's user interface is built using X windows, the application can be used remotely over a network from any X server terminal. In that event, performance of JEDMICS/CADA will degrade during visual QA sessions since large amounts of graphical data must be sent across the network which may have high traffic or low bandwidth.

5.3.2 Hardware

A standard Ethernet Network Interface Card (NIC) with the appropriate drivers necessary for the OS to support the communications link would be required for JEDMICS/CADA to take advantage of network resources for greater functionality and integration with current QA processes. All Sun SPARCstation compatibles come with an Ethernet NIC and are network ready.

5.3.3 Software

Support of TCP/IP and NFS protocols are required for JEDMICS/CADA to take advantage of network resources. All Sun SPARCstation compatibles come with the appropriate X windows client, TCP/IP.

5.4 Interfaces

JEDMICS/CADA generates PostScript files for reports and raster images which are spooled to the printer using Unix's standard lp command.

5.5 Summary of Impacts

Since JEDMICS/CADA is a stand-alone system geared towards enhancing QA of incoming data there are no major impacts to the input or output of current repository systems. JEDMICS/CADA serves primarily as a "first-stop" QA station where data would be reviewed before being input into the repository system. For more information on how JEDMICS/CADA impacts QA personnel, refer to Section 2.4.2.

5.5.1 Automated Data Processing (ADP) Organizational Impacts

There are no Automated Data Processing (ADP) organizational impacts as a result of JEDMICS/CADA's use in the QA process. There would not be a modification, addition, or elimination of positional responsibilities. QA operators will need to undergo training to use the JEDMICS/CADA software. Training sessions will specify the strengths and differences JEDMICS/CADA imposes on the typical QA process.

5.5.2 ADP Operational Impacts

There are no major ADP operational impacts as a result of JEDMICS/CADA's use in the QA process. JEDMICS/CADA would be used as a "first-stop" QA station where data would be reviewed before being input into the repository system. Currently, incoming data are loaded into a temporary location on the repository system and reviewed by a QA operator at a GDT station. Accepted documents are written to the optical disk of the repository system and a rejected documents report is sent back to the originating source.

Using JEDMICS/CADA, incoming data would be reviewed at the JEDMICS/CADA workstation then the QA operator would forward the accepted documents to the repository system for a batch accept.

5.5.3 ADP Development Impacts

CADA has already undergone extensive testing during a seven month alpha test period (refer to Final Report: CADA Field Tests and Productization, 18 October 1994). Various features were finalized and several bugs/performance problems were addressed with the help of field test users at the Army (CECOM/MICOM), Air Force (Tinker AFB), and Navy (NADSURWARCEN Ordsta). As a result, software problems should be primarily limited to learning curve/configuration issues for new users.

5.6 Failure Contingencies

JEDMICS/CADA displays various messages throughout the application when errors occur. Most of the messages are self-explanatory and non-fatal. The user will be able to deal with this situation which usually involves bad data entry on the user's part. Refer to Appendix E in the JEDMICS/CADA End User's Manual for a listing of how to deal with basic JEDMICS/CADA error messages. Any errors that are serious in nature are specified and the system administrator should be called to handle this situation. Refer to the JEDMICS/CADA Maintenance and Computer Operator's manuals for instructions to deal with fatal situations.

5.7 Assumptions and Constraints

The following assumptions and constraints are present when using the JEDMICS/CADA tools:

- Content QA is still required to be performed by engineering personnel. JEDMICS/CADA does not check the content of an image except for the ID data located in the title and revision blocks of a drawing.
- JEDMICS/CADA performs processing on only one batch at a time.
- The QA personnel are responsible for releasing processed batches from CADA to pending storage.
- JEDMICS/CADA timing performance is very dependent on the processing power of the CPU, amount of RAM present, system load, and content of data used. Expect performance to vary under different conditions.

SECTION 6 SECURITY

6.1 Background Information

The JEDMICS/CADA software has no security classification. Once approved for use by the Government, full documentation and software code (with the exception of third party COTS products) will be available.

6.2 Control Points, Vulnerability, and Safeguards

Security for this release is limited to the password a user has for his or her Unix account on the JEDMICS/CADA workstation. Once a user has successfully logged in, JEDMICS/CADA does not provide any additional identification or security to stop users from retrieving batches from pending or permanent storage, viewing images, overriding evaluation results, modifying index information, or releasing batches to pending storage. The system administrator should be responsible for controlling who has access to the system and training users on basic security precautions when using a multi-user computer operation system.

6.2.1 Control Points

There are various control points that are vulnerable security points since the software does not have a requirement for any individual identification or security to stop users from performing the following activities from within JEDMICS/CADA once logged into an account:

- retrieving batches from Pending or Permanent,
- viewing images,
- overriding automated evaluation results,
- modifying index information, and
- or releasing batches to Pending.

6.2.2 Vulnerability

Once image data are loaded into a temporary location specified by JEDMICS/CADA's configuration, a user with the proper privileges could view, copy, remove, or modify the files.

6.2.3 Safeguards

The system administrator/security officer determines user privileges when the user account is established and also has the ability to change established privileges. The system administrator should set each user's account with his or her own JEDMICS/CADA directory structure that can not be written or read by any other user. This, coupled with the user account passwords, should provide enough basic protection from users manipulating anybody's active data.

An auto-log out time-out should be set on user accounts to avoid unauthorized use when users inadvertently leave their JEDMICS/CADA sessions unattended for long periods of time. Users

should also change their passwords on a regular basis and the passwords should not contain information that can easily be guessed (e.g., phone numbers, addresses, family names, etc.).

6.3 System Monitoring and Auditing

JEDMICS/CADA logs information for each JEDMICS/CADA work session. The actions performed, along with the date and time stamp, are written to log files. The files have information on operator actions as well as evaluation parameters used and JEDMICS/CADA evaluation results. The log files are written as events occur in order to avoid loss of information.

6.3.1 Journalizing

Figure 3 shows a sample log file *(cada.log)*. The *cada.log* file logs events related to basic operations such as loading, evaluation statistics, etc.

Need copy for this

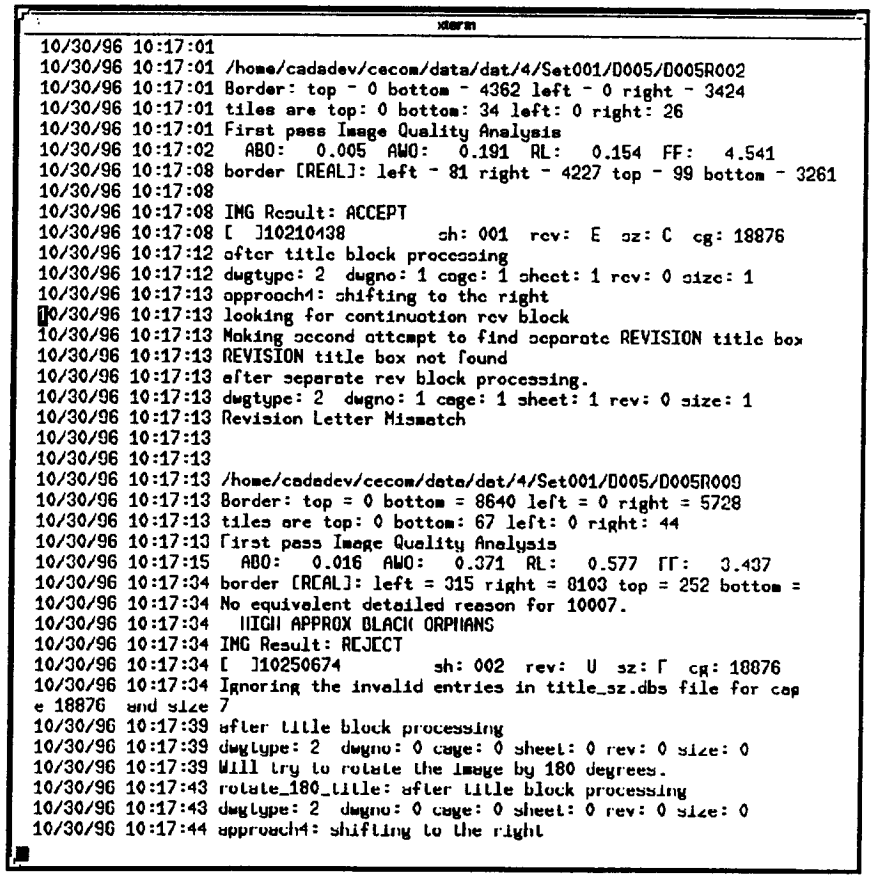

Figure 4. CADA Session Event Log

JEDMICS/CADA also maintains log files specific to each loaded batch. These log files indicate the operator actions performed on each batch and the time between events. The naming convention is *batch-batch_number.log*.

At the start of a new JEDMICS/CADA session, the previous *cada.log* file is archived in a feedback directory. The naming convention used for the file name is as follows: ***year.month.day-hour.minute.log***. The log details events up to the time specified by the filename.

6.3.2 Audit Trail

The user is not required to initiate any actions to start an audit trail. All logging of user activities is done automatically. However, the same concerns of security apply to these log files. A user with the proper privileges could view, copy, remove, or modify the log files. As stated before, the system administrator should set each user's account with his or her own JEDMICS/CADA directory structure that can not be written or read by any other user.

APPENDIX A
Commercial-Off-The-Shelf (COTS) Products

In order to use the JEDMICS/CADA executable software package, it is required that run-time licenses be obtained for use of the COTS products. It is unlawful to use this package without obtaining the licenses. The Government does not endorse these COTS products and they may be used at the discretion of the user. Other equivalent COTS products may be substituted upon obtaining the source code from the Government. The COTS products that are incorporated into the JEDMICS/CADA executable software are shown below. The licenses may be obtained from IC&G Systems, the OEM distributor, for these products for use in JEDMICS/CADA or directly from the vendors.

APPLICATION	PRODUCT	VENDOR
• OCR/ICR Software	NestorReader™ Part No. NR-CADA	NCS™ Recognition Products One Richmond Square Providence, RI 02906 Attn: David P. Wright
• Imaging Library 1 ScanFix™	TMM Sequoia Part No. CADA-SF	
• Imaging Library 2	UniSoft Imaging Library Part No. USCADA-V6	Unisoft Imaging 4606 N. Britton Road Stillwater, OK 74075 Attn: Calvin Aiken

Source for Licenses: Integrated Computer & Graphics Systems
809 East Redwood Court
Highlands Ranch, CO 80126
(303) 470-7262 (Tel. or FAX)

Disclaimer: The use of these COTS products does not constitute an endorsement or approval of their use. This software may not be cited for the purpose of advertisement. Any of these products may be replaced by equivalent products due to the availability of the CADA source code.

ADDENDUM
to
Computer-Assisted Data Acceptance (CADA) Functional Description dated 28 November 1995

This addendum identifies changes to the Function Description dated 30 November 1995 that was prepared during the early development of the JEDMICS/CADA System.

A. Deletions and Changes

1.2.2 Contractor Deliverables

Title	Date
CADA End Users Manual	*28 October 1996*
JEDMICS/CADA Contractor	*28 October 1996*
Final Report: Application of CADA to JEDMICS	*28 October 1996*

Paragraph 3 and part of last sentence of 2.2 Objectives

The CADA system will provide JEDMICS QA personnel with the ability to perform automated QA according to the following two scenarios:

> **PENDING BATCHES** - The user is able to perform image quality and index information verification for a specified batch which is resident in JEDMICS Pending. Using CADA, the user can change index information and QA flags which will subsequently be changed within the given batch in Pending.
>
> ~~**LARGE FORMAT DRAWINGS FROM PERMANENT STORAGE** - The user is able to initiate index information corrections for large format drawings which are contained within JEDMICS Permanent.~~

The objective of CADA, through automation, is to cost-effectively perform 100% QA for batches resident within JEDMICS Pending ~~and correct inconsistencies for large format drawing header information within JEDMICS Permanent~~

In Paragraph 5 of 2.4Proposed Methods and Procedures for the Application of CADA Techniques to JEDMICS

If the user requested a batch from Permanent then the batch will only undergo the image conversion, a subset of the index verification ~~(to check for inconsistencies with large format drawing)~~, visual QA, report generation, and batch output.

In Paragraph 3 of 2.4.1 Summary of Improvements

JEDMICS/CADA also provides the capability to automate the analysis of index information within a drawing and compare it with the corresponding index information used to store the image. Again, the operator has the opportunity to analyze the JEDMICS/CADA results and override them if he or she chooses to do so~~. These index checks will be expanded to check for inconsistencies in index information for large format drawing which reside in Permanent.~~

Paragraph 4 of 2.4.2.2 User Operational Impacts

~~CADA will also be used to correct index inconsistencies found within Permanent for large format drawings. These inconsistencies will be detected and corrected by JEDMICS/CADA. QA personnel will then direct JEDMICS/CADA to submit any changes to Pending.~~

3 Paragraphs in 2.5.1 Assumptions

- ~~The index information which will be updated for large format drawings is limited to the drawing size.~~

- ~~Large format drawings are correctly indexed within JEDMICS.~~

- ~~Drawing revision extraction is limited to C, D, E J, G, and H size (multiframe) drawings.~~

1 Paragraph in 2.5.2 Constraints

PERMANENT STORAGE OPERATIONS

- ~~The upload operation will utilize Pending and will therefore, re-insert only the index header information.~~

Part of sentence in last paragraph in 3.1.1 Accuracy and Validity

This system will address the need of ensuring good quality image ~~files and the need to correct large image format index information~~. Due to automation, it will help to reduce the time factor associated with such jobs.

A portion of 3.2 Functional Area System Functions

Batch Submittal: This function is responsible for the submittal information from CADA to JEDMICS. All submittals will be to Pending, however, depending on the batch processing one of two scenarios will take place.

If the batch was originally retrieved from Pending, then all QA flags, as described in section 4.2.3 may be set. The existing batch will be updated accordingly based on the changes made from the JEDMICS/CADA workstation. ~~If the batch was originally retrieved from Permanent then a new batch will be created which consists of only the index information which must be corrected. They will be submitted as new and the incorrect index information will be marked for deletion.~~

~~**Evaluate Index Information:** This function is responsible for performing automated evaluation of all index information. The accuracy of the index information will be compared against the same information which is present on the image face. The user will only have to initiate the evaluation. Once started, the evaluation will run unattended.~~

~~This function is available to be run only against batches downloaded from the Pending data base.~~

Last line of first paragraph and part of last paragraph of 3.3 Inputs and Outputs

The JEDMICS/CADA system will accept the input of JEDMICS C4 image files and the corresponding index information for each of these files. All inputs will be addressed within JEDMICS/CADA as a batch. If the batch is received from JEDMICS Pending, JEDMICS/CADA will retain the batch number for consistency but will assign its own batch number for internal processing. ~~If the batch is received from Permanent, JEDMICS/CADA will assign a unique batch number for processing within JEDMICS/CADA~~.

For batches which originated from the Pending data base, the above information will be updated for the original batch. ~~For batches which originated from Permanent, the above information will be supplied in a newly created batch within Pending~~. The changed index information will be added as a new drawing. The old information will be tagged for deletion.

Sections of 4.2.2 Batch Submittal

JEDMICS/CADA will use the existing JEDMICS API to submit batches to the Pending data base. The manner in which this is done and the information which JEDMICS/CADA will change is based on the original location of the batch (Pending ~~or Permanent~~).

Due to restrictions of the API, Pending batch information cannot be modified. Therefore, when processing Pending batches, JEDMICS/CADA will create a new batch with all of the original information, plus the changes approved from automated QA at the JEDMICS/CADA workstation. ~~When processing Large Format Drawings, JEDMICS/CADA will create a new batch which will consist of the changed index information which will be added to the permanent storage along with the original index information which will be tagged as deleted from Permanent.~~

Section in 4.2.3 Input/Output

The input of batches downloaded from JEDMICS occurs strictly for allocating and populating internal data structures required by JEDMICS/CADA. Upon input, the index information will be appropriately stored and the batch ID will be established. This ID will be retained from pending storage ~~or assigned if the batch is received from Permanent~~.

The output of batches which have been processed through JEDMICS/CADA involves establishing the appropriate set to be submitted to Pending. The output set will consist of only the index information, as the image files are not changed by JEDMICS/CADA. This index information will be appropriately set based on the operations performed within JEDMICS/CADA.

The output will also involve the mapping of JEDMICS/CADA results to QAFlags understood by JEDMICS. The following table shows this mapping.

Image Quality Errors

JEDMICS/CADA Error	JEDMICS/CADA Error Description	JEDMICS QA Flag
DECOMP_ERR	Detected when JEDMICS/CADA has had a problem decompressing the image.	JMX_POOR_COMP
HIGH_BLK_ORPHS	This is the ratio of black run lengths of one in non-vertical mode to the total number of white pixels. It approximately indicates the black orphan pixels within the image.	JMX_TOO_DARK
HIGH_WHT_ORPHS	This is the ratio of white run lengths of one in any mode to the total number of black pixels. It approximately indicates the white orphans pixels within the image.	JMX_TOO_LIGHT
HIGH_FF	This technique takes the fill factor and compression ratio of the compressed image into consideration. The assumption is that a noisy image (high fill factor) is compressed to a lesser degree (low compression ratio).	JMX_TOO_DARK
LOW_FF	This technique takes the fill factor and compression ratio of the compressed image into consideration. The assumption is that a sparse image (low fill factor) is highly compressed (high compression ratio).	JMS_TOO_LIGHT
HIGH_RLRAT	Run length analysis is a means of detecting noise data in compressed image	JMX_TOO_DARK

JEDMICS/CADA Error	JEDMICS/CADA Error Description	JEDMICS QA Flag
	data. A high ratio indicates a noisy image. A low ratio indicates a noise-free image.	
HIGH_PEAK_TILE	Technique used to detect images that are faded in small areas. Image is divided into 256 x 256 pixel tiles and the noise level is detected for each tile.	JMX_TOO_LIGHT

If the batch was originally received from Pending and image quality and index verification were performed then a new batch would be generated which would consist of all updated information from within JEDMICS/CADA as well as all unchanged information retained by JEDMICS/CADA.

~~If the batch was originally received from permanent storage then the updated index information would be submitted as new and the index information it is correcting would be tagged to be deleted.~~

Last Paragraph in **4.2.5 Evaluate Index Information**

~~In addition to this checking, all images which are candidates for large format drawing (LFD) index verification will be checked accordingly. This checking is discussed in Section 4.2.6.~~

~~All of 4.2.6Evaluate Large Format Drawing (LFD) Index~~

~~Images will be considered candidates for large format drawing processing if they are of size C, D, E, J, G or H. The physical size of the image will be cross-verified with the size contained in the corresponding index information. The large format drawings are split across multiple frames. These multiple frames have the identical index information (as per their six ID fields mentioned in the last section), making them candidates for large format drawings. These multiple frames will then be evaluated for additional characteristics like continuation title block and a determination made as to its large format size. The size field in the index information will then be updated, if necessary.~~

B: Contract Additions

1. Install a JEDMICS/CADA Prototype at JEDMICS CECOM for automated QA and index validation of migrated data to Permanent. Training will also be provided to the operator.

2. Install JEDMICS/CADA Prototypes at JEDMICS/Warner Robins, JEDMICS/MICOM, and JEDMICS/Port Hueneme and provide operator training at these sites.

3. Incorporate Revision Block Zone - A configurable option allows users to turn off strict revision information checking. This option, when used, will ignore revision zone information contained in the separate revision block of the image. It will, instead, only match the revision letter to the revision letter contained in the index information.

4. Incorporate Rejects by Zones Reporting - The CADA Evaluation Report will include missing borders and quality rejects by zones within the image area.

5. C4 Output Report - The CADA Evaluation Report will identify specific C4 format problems such as corrupted or irregular header structure, and bad compression data.

6. Image Orientation/Rotation - While evaluating C4 images from Warner Robins, it was found that some of the images may be rotated in any direction. The C4 header does not presently provide orientation and rotation information as is found in CALS Type 1and Type 2 headers.
7. This caused JEDMICS/CADA ID failures due to the inability to find the title block. It became necessary to develop the capability to locate the correct orientation of the image and include this in the JEDMICS/CADA System. This added design task was authorized by the JEDMICS program office.

8. Rights Block - This option would allow the user to ignore image and quality failure if it occurs only within the Rights Block area (upper left corner) of an image. This option was not incorporated since the orientation/rotation issue was of a higher priority due to the images found at Warner Robins that had been rotated.

9. Requirements Papers - Three requirements papers were prepared at the request of the JEDMICS Program Office: Drawing Title Recognition, Rejects to Pending Data Base as Separate Batches, and Output Rejects to Pending as Separate Batches.

10. ISDN Capability - This capability was added to allow higher throughput rates when transferring large batches of image data via the Internet. The delays within the numerous Government operations from the source JEDMICS site to ACCURATE has not improved the throughput to the extent desired.

C. Other

In addition to the contract additions identified other improvements and methods of processing were required as the design progressed and more information was obtained on the JEDMICS API. The design began with API 2.0 and has progressed until API 2.5 was incorporated. These API improvements required changes in the way CADA accessed JEDMICS and other restrictions such as the orientation, batch size within Pending that limits the "query hit limit" to 1000, and the unique index of Pending. All this required changes in the way we originally planned to access and process the batches of images within Pending.

Extensive testing of the images obtained from the sites to improve accuracy of QA and ID recognition was done. The access and image transfer time posed problems in that we could not achieve the testing of JEDMICS "live" image data on the scale that we had planned.